

POU92002 0018 US1  
1/30

KNOP ET AL

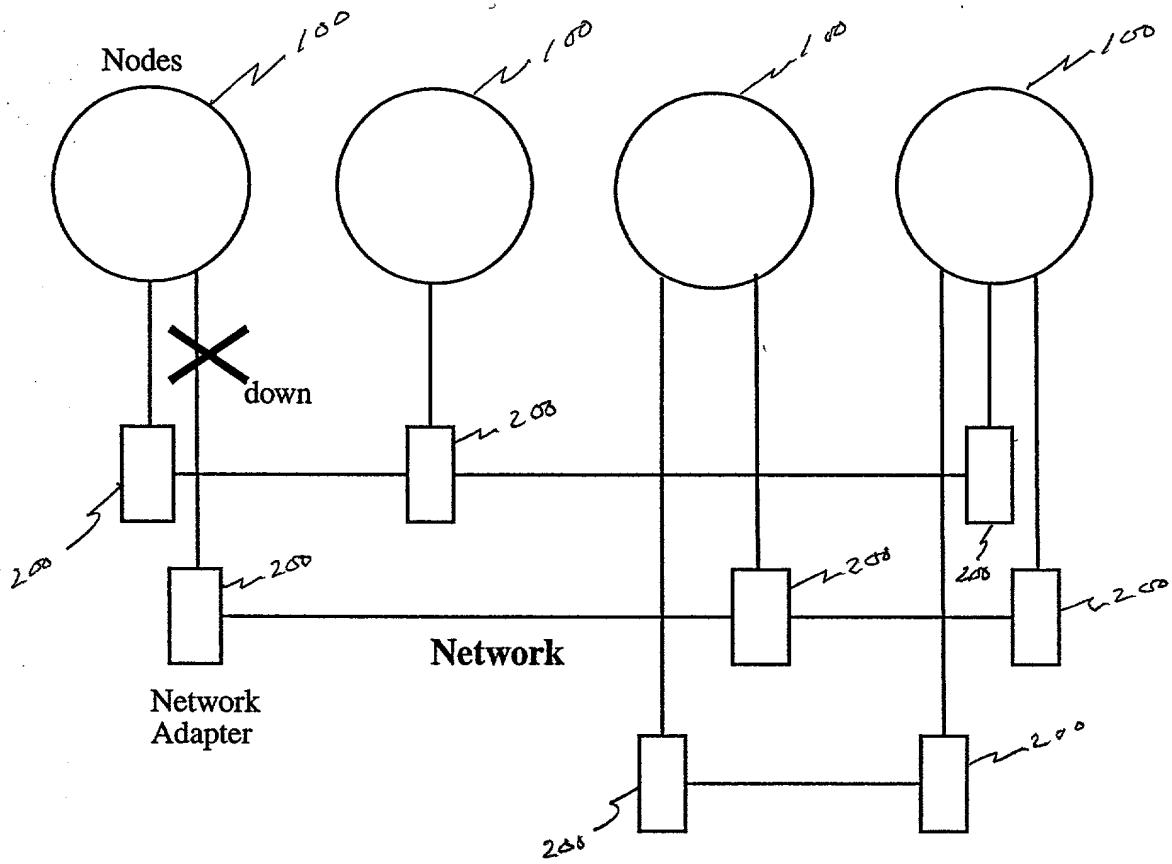


Figure 1: Nodes and physical connectivity

2/30

PO# 920020018451

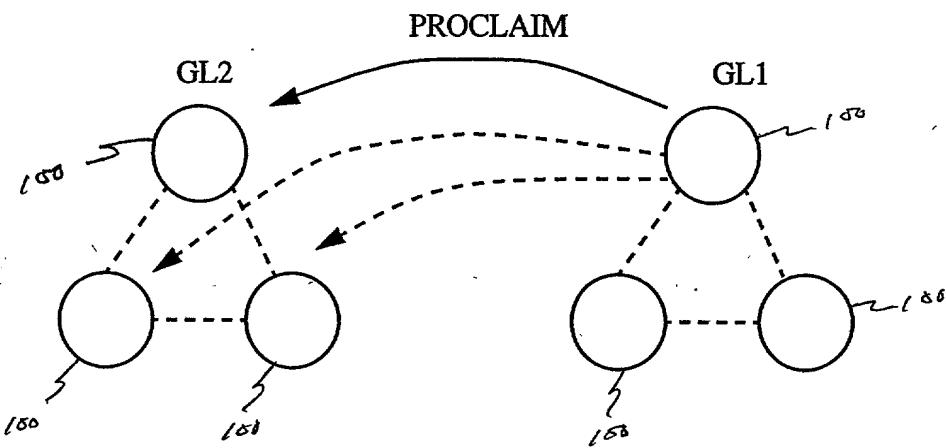


Figure 2a) JOIN Protocol: PROCLAIM message

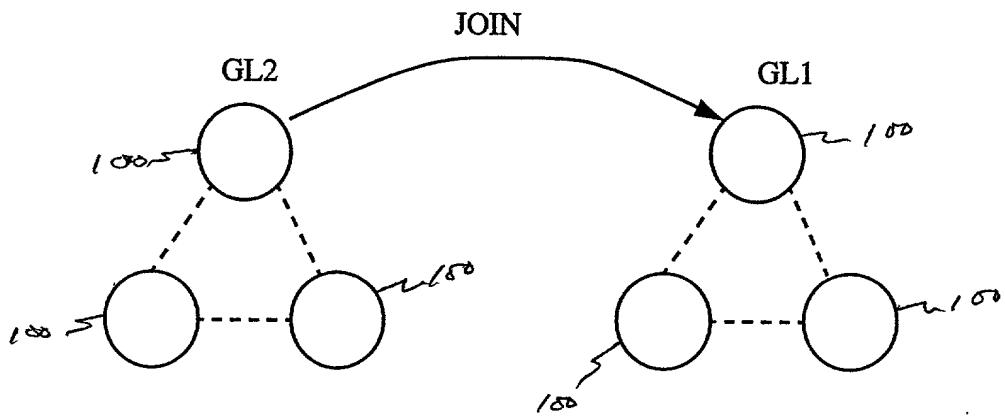


Figure 2b) JOIN Protocol: JOIN message

3/30

Pou 920020018451

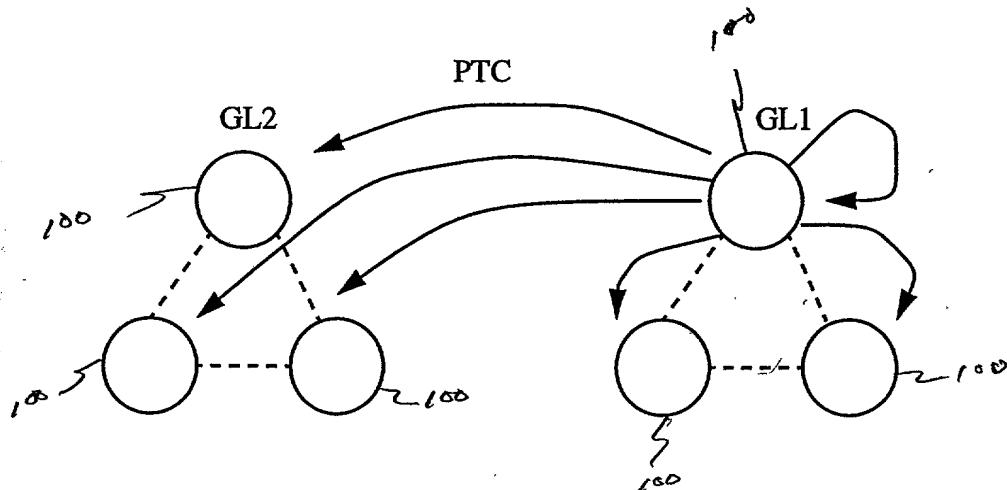


Figure 2c) JOIN Protocol: PTC message

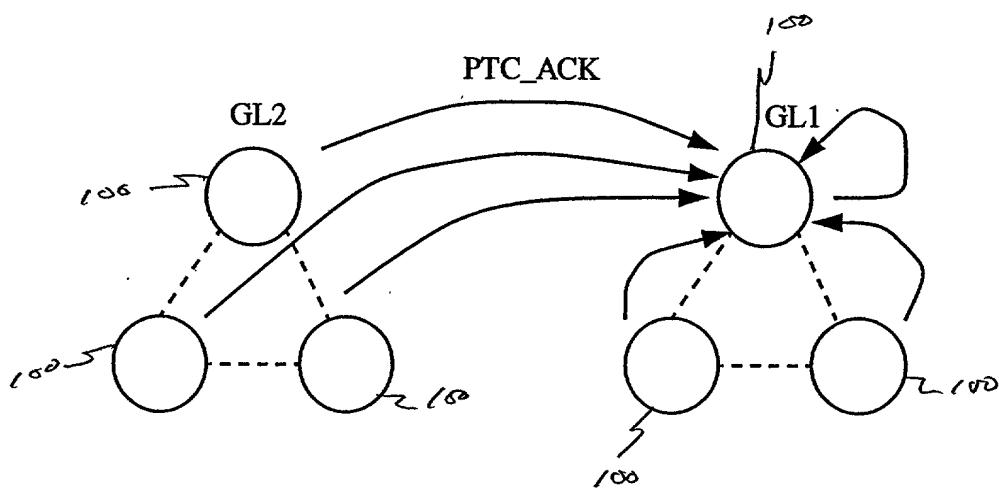


Figure 2d) JOIN Protocol: PTC\_ACK message

4/30

Pou 920020018 US1

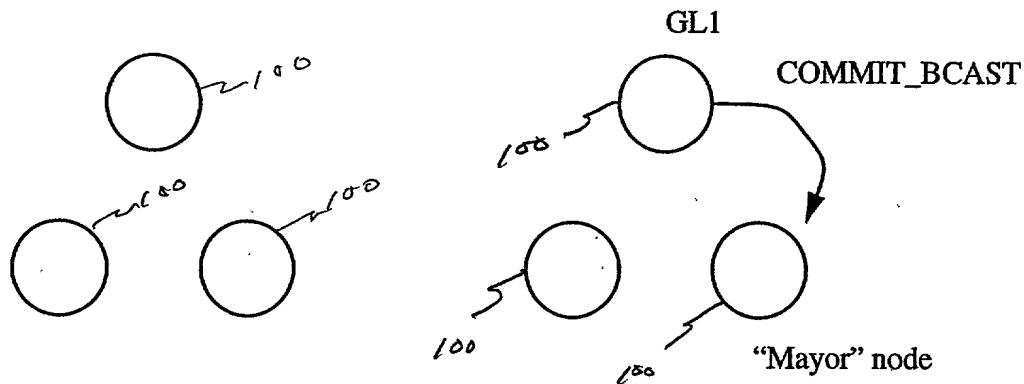


Figure 2e) JOIN Protocol: COMMIT\_BCAST message

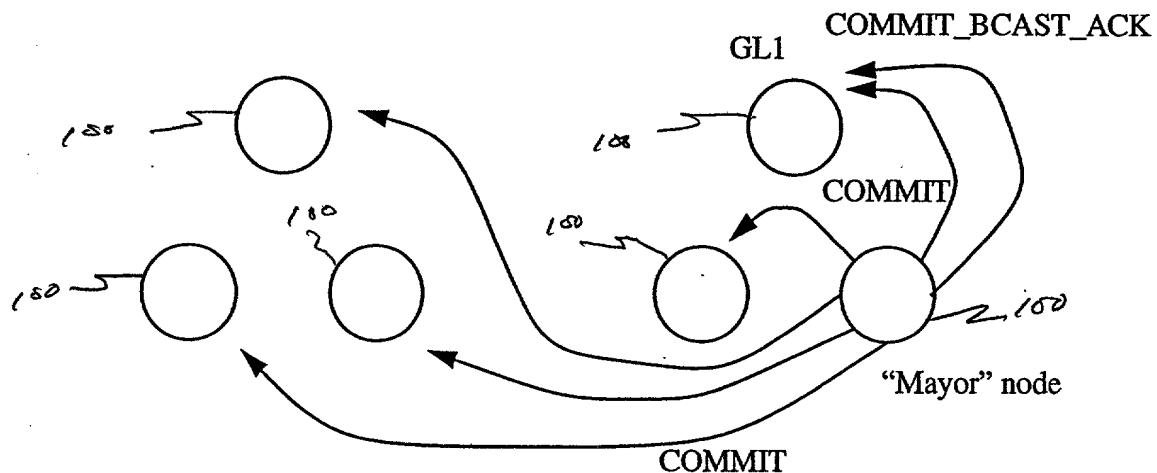


Figure 2f) JOIN Protocol: COMMIT and COMMIT\_BCAST\_ACK messages

5/30

POA 920020018451

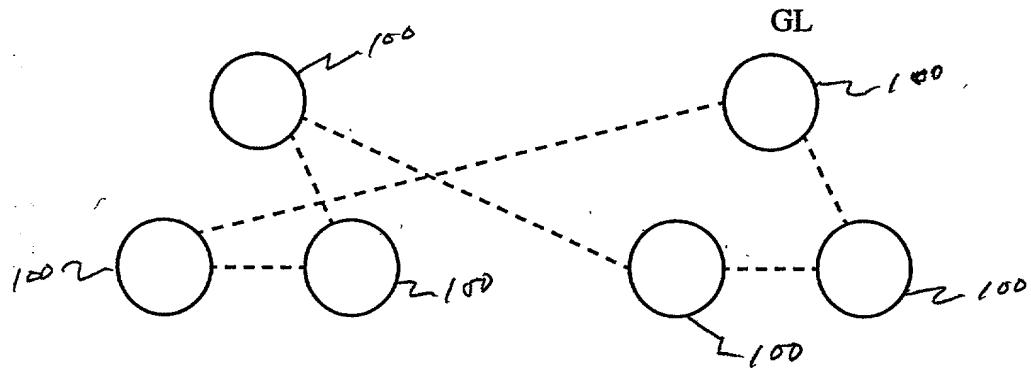


Figure 2g) JOIN Protocol: new group formed after completion of protocol

Figure 2: JOIN protocol

6/30

Pou 920020018451

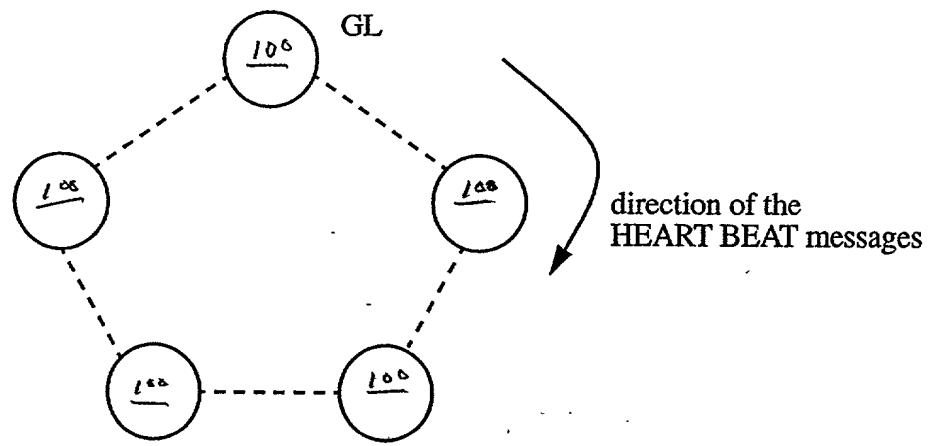


Figure 3a) DEATH Protocol: initial state: heartbeat ring

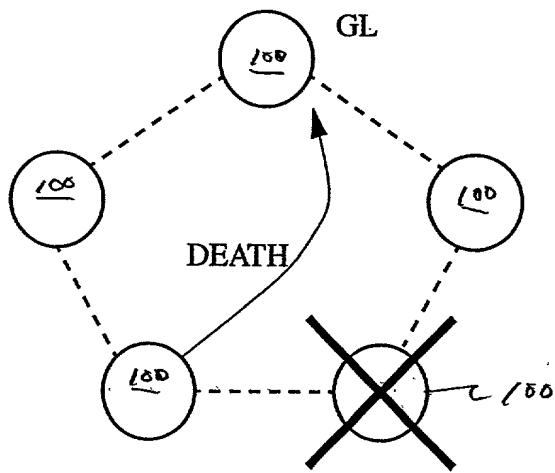


Figure 3b) DEATH Protocol: DEATH message

7/30

Pub 92002 0018 us1

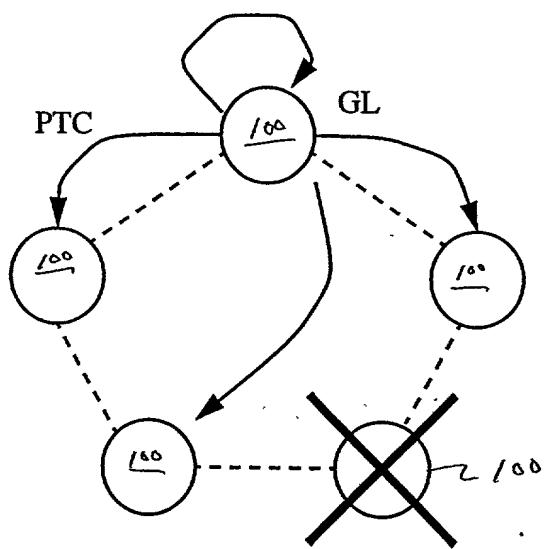


Figure 3c) DEATH Protocol: PTC message

Figure 3: DEATH protocol

8/30  
Pou920020018451

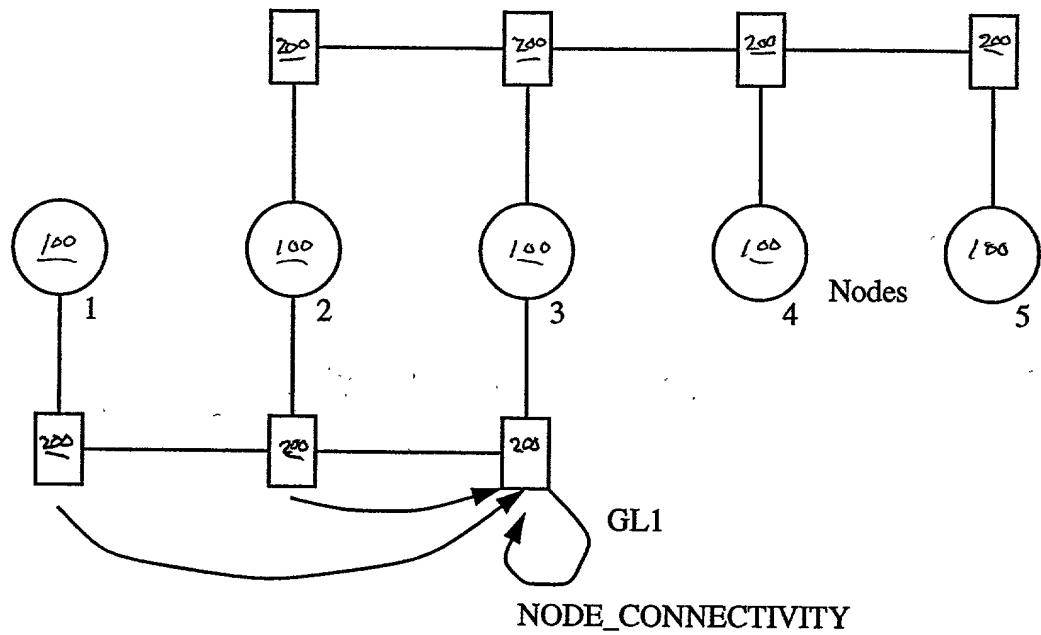


Figure 4a) Node Reachability Protocol: NODE\_CONNECTIVITY message

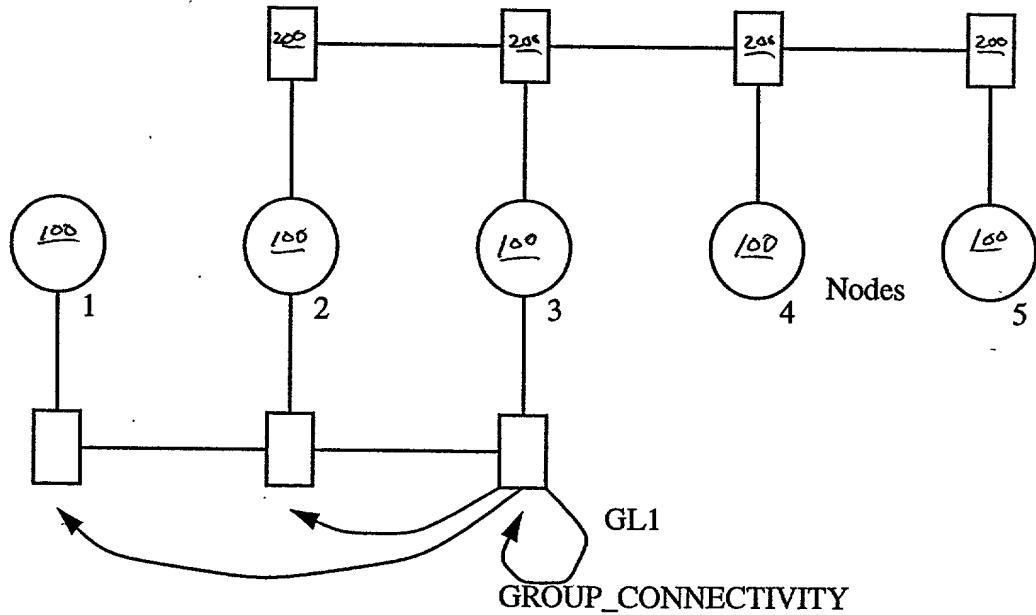


Figure 4b) Node Reachability Protocol: GROUP\_CONNECTIVITY message

9/30

P04920020018 US1

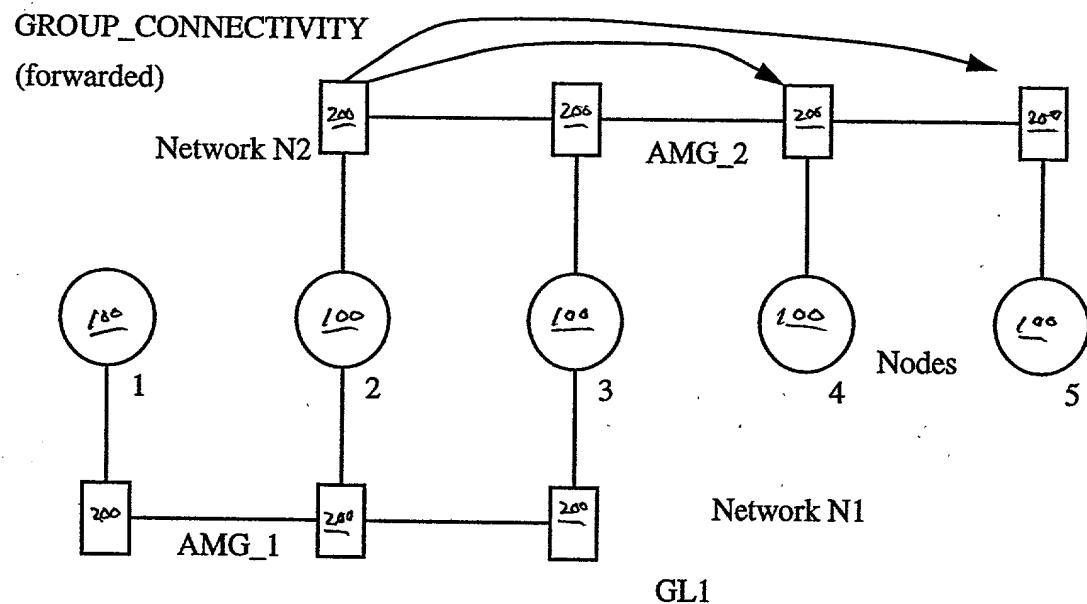
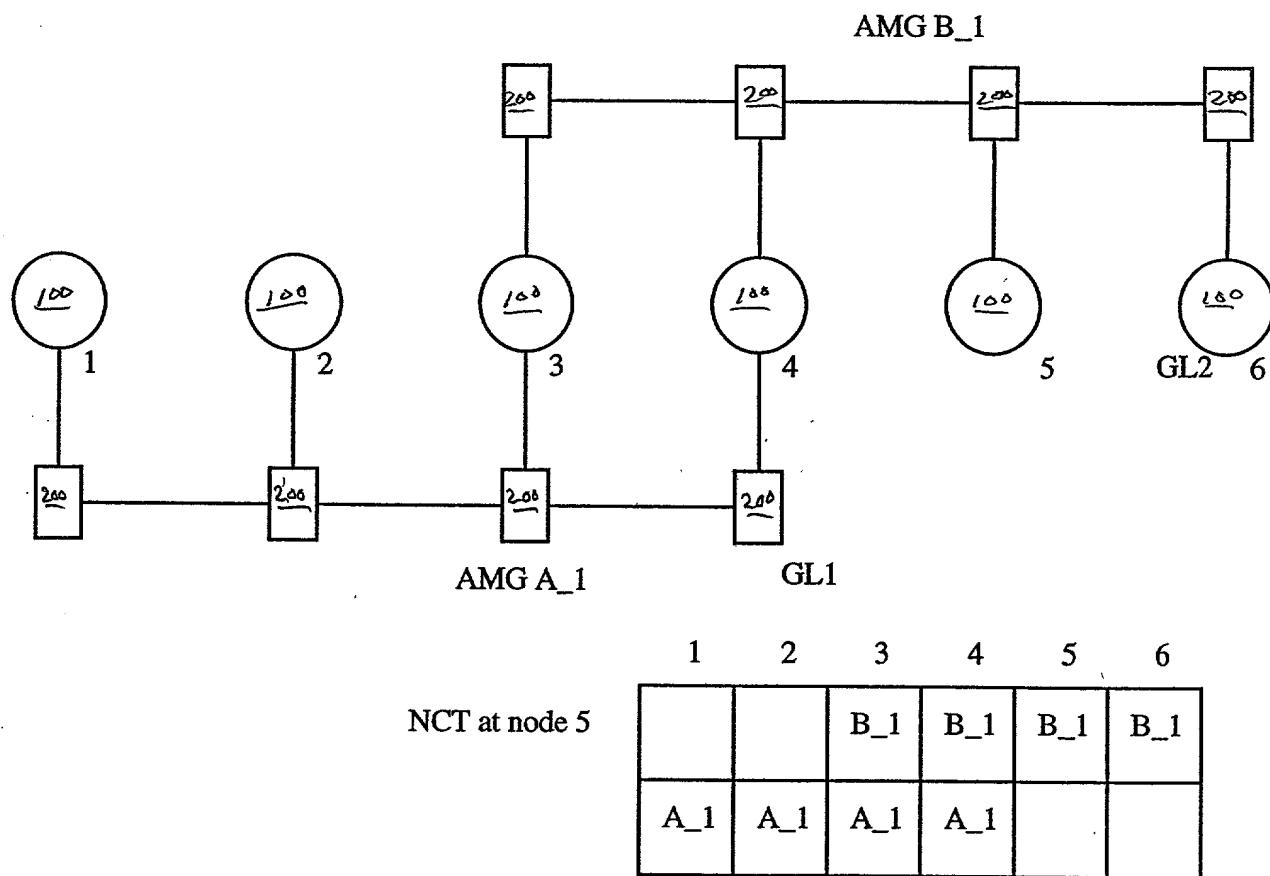


Figure 4c) Node Reachability Protocol: forwarding of GROUP\_CONNECTIVITY message

Figure 4: Node reachability protocol: NODE\_CONNECTIVITY and GROUP\_CONNECTIVITY messages

10/30

P04920020018451



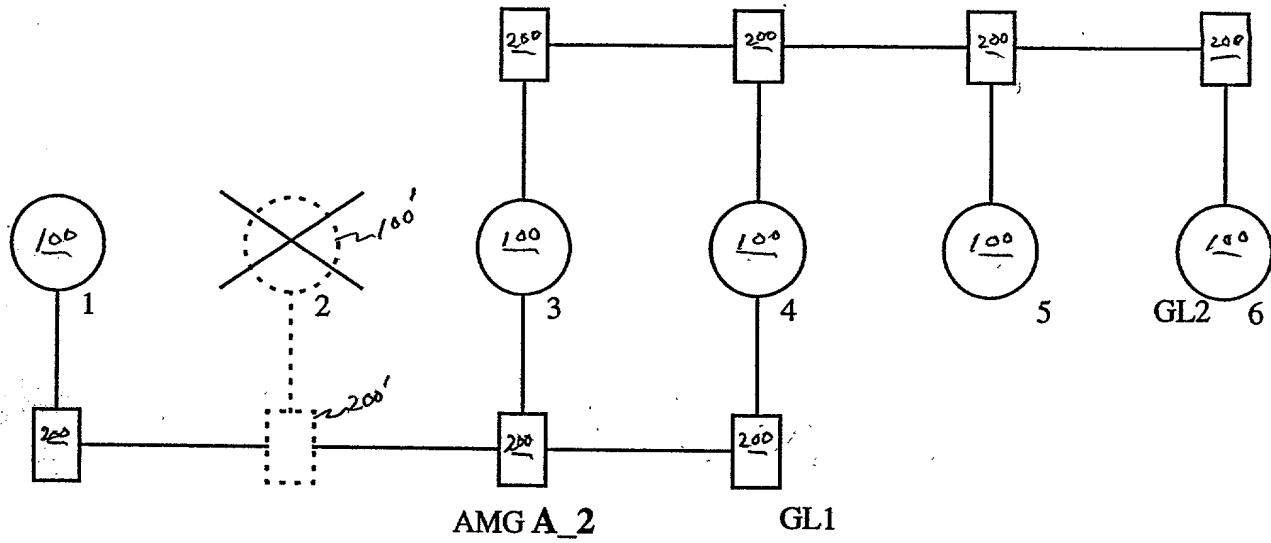
5a) Initial situation

Figure 5: Topology Propagation Scenario: node death

11/30

P04920020018 US1

AMG B\_1



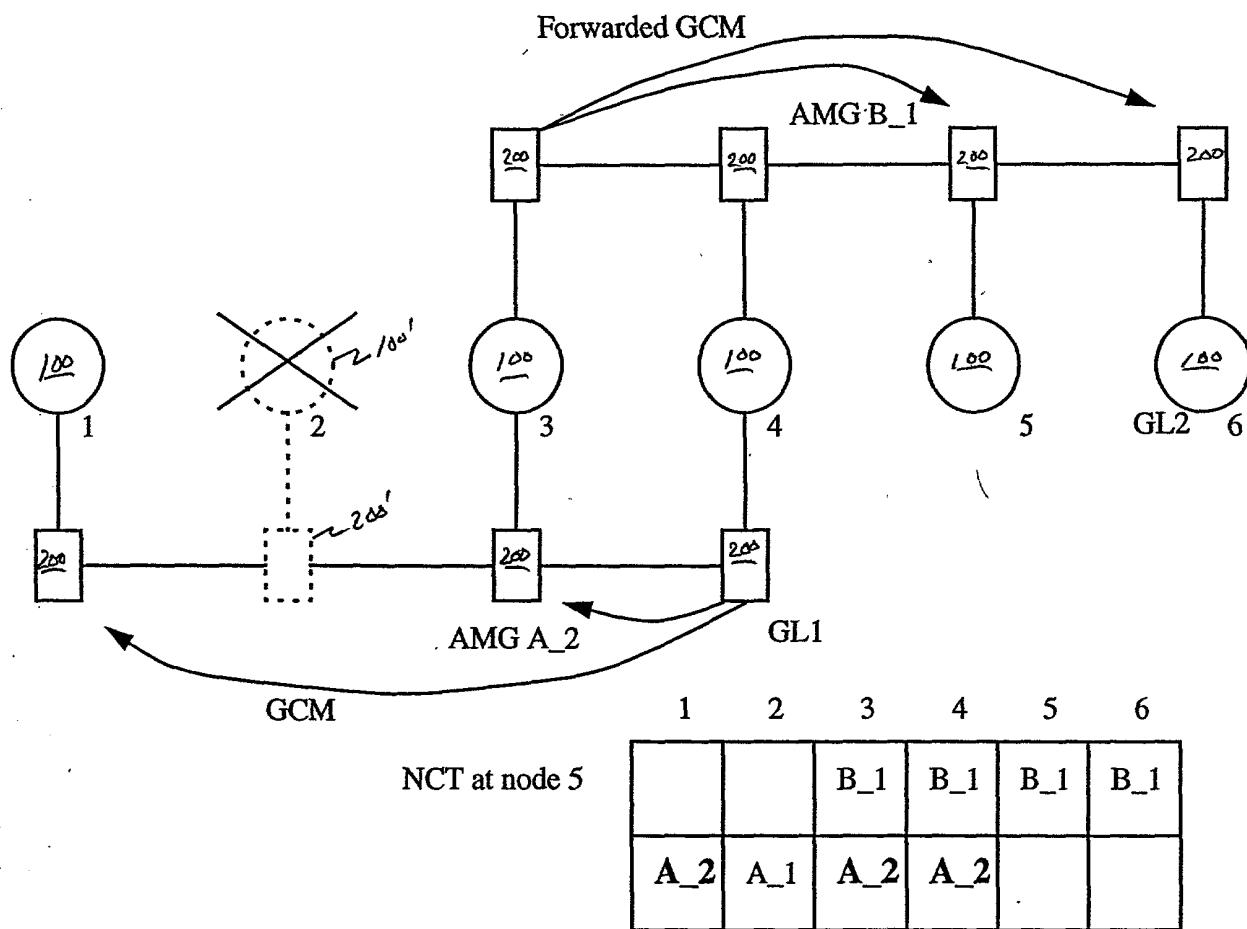
NCT at node 5

		B_1	B_1	B_1	B_1
A_1	A_1	A_1	A_1		

5b) Node 2 dies: Nodes 1, 3, and 4 form AMG A\_2

12/30

Pou920020018451



5c) GCM for AMG A\_2 is propagated to all nodes

13/30

Post 920020018 us1

Network A: 10 seconds detection time  
Network B: 40 seconds detection time

Nodes 1 and 2 are forming AMGs on networks A and B

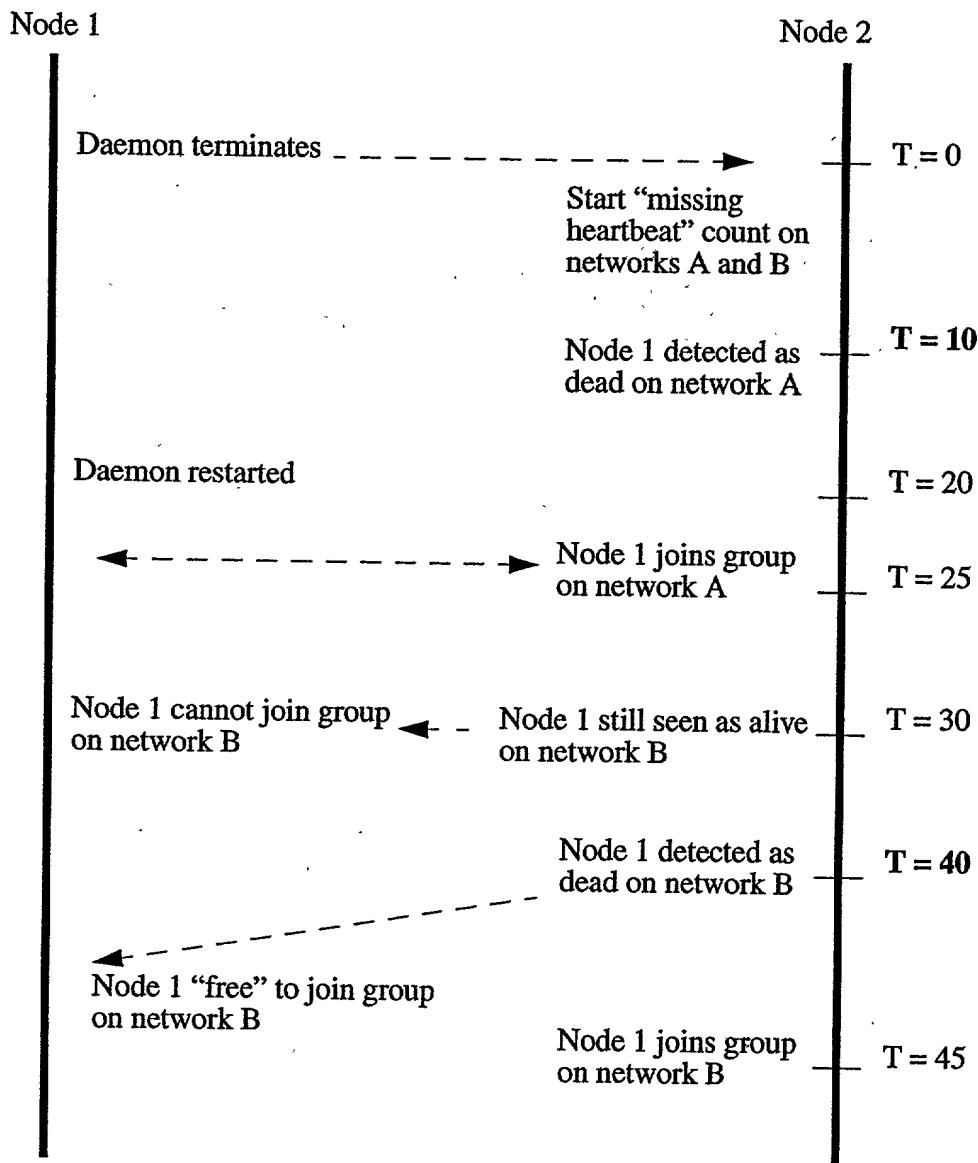


Figure 7) Inconsistency caused by quick daemon restart in the presence of different detection times for each network: the daemon on node 1 goes down and is restarted, but this is never detected by node 2.

14/30

Post 920020018 us1

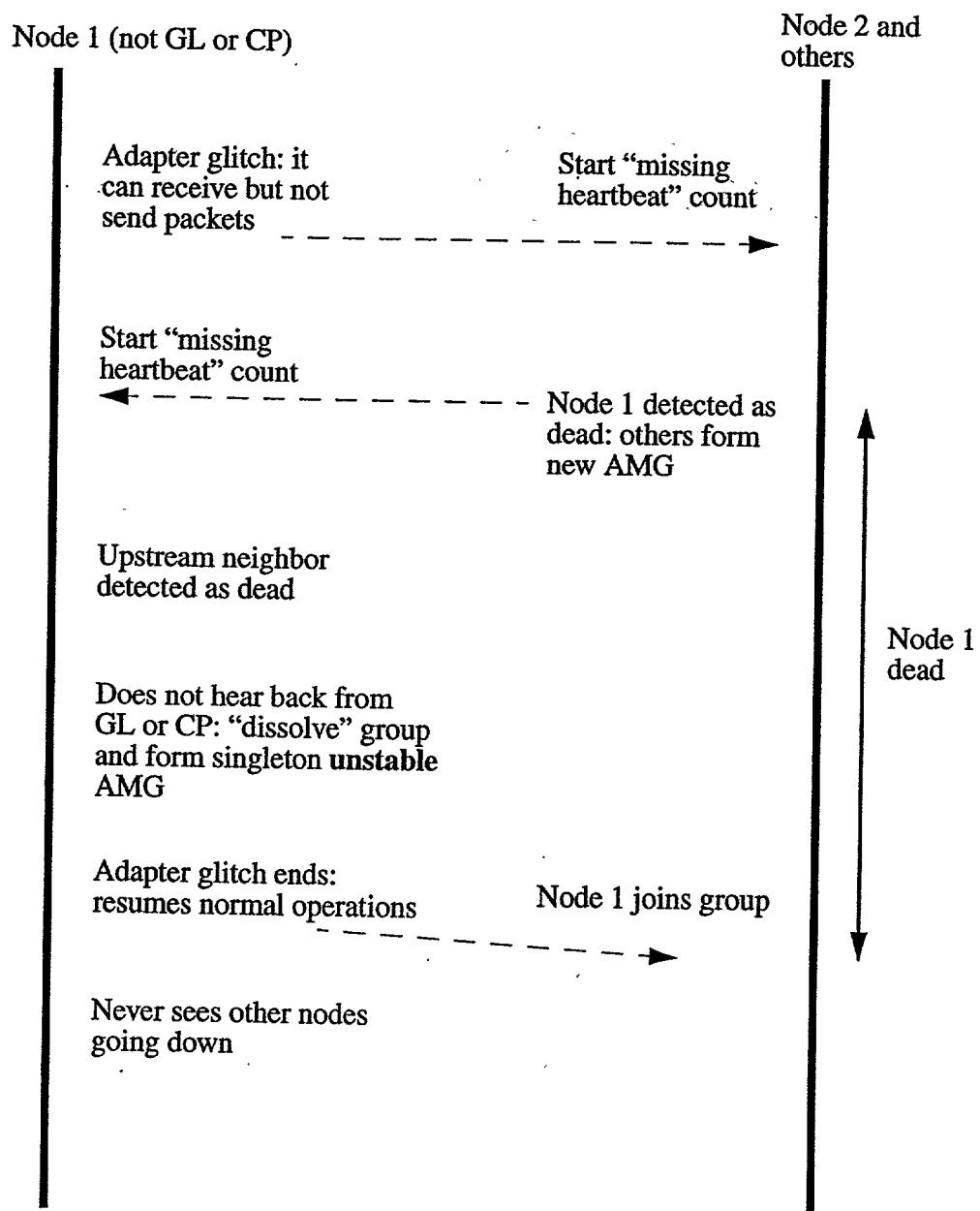


Figure 7) Inconsistency caused by temporary communication problem in an adapter: node 1 is not the GL or CP in its group. Node 1 never notices other nodes as down, though the others do see node 1 being unreachable.

15/30

Post 920020018 251

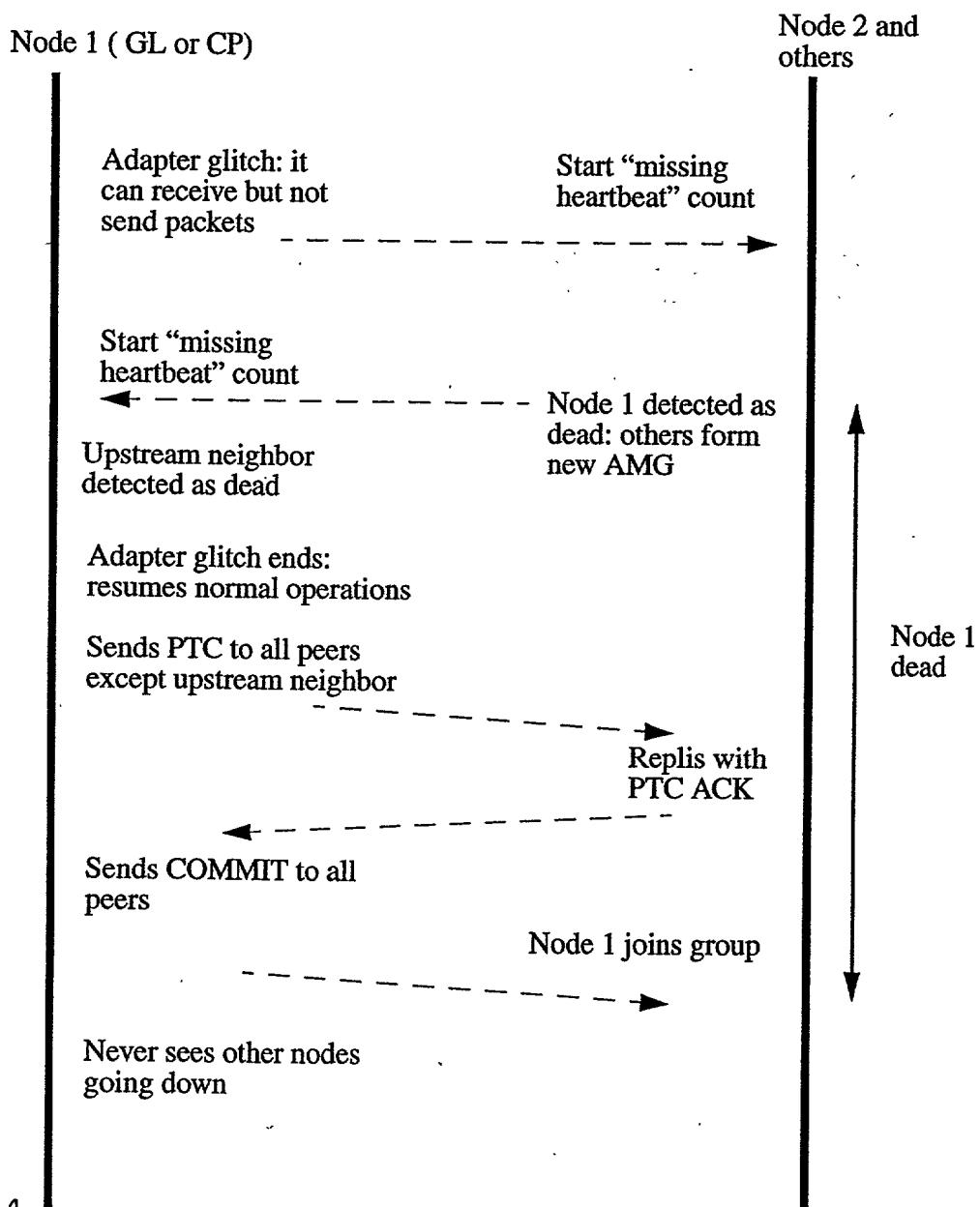


Figure 6) Inconsistency caused by temporary communication problem in an adapter: node 1 is the GL or CP in its group. Node 1 never sees the other nodes as unreachable, while the others do see node 1 as unreachable for a period.

16/30

P0492002 0018 451

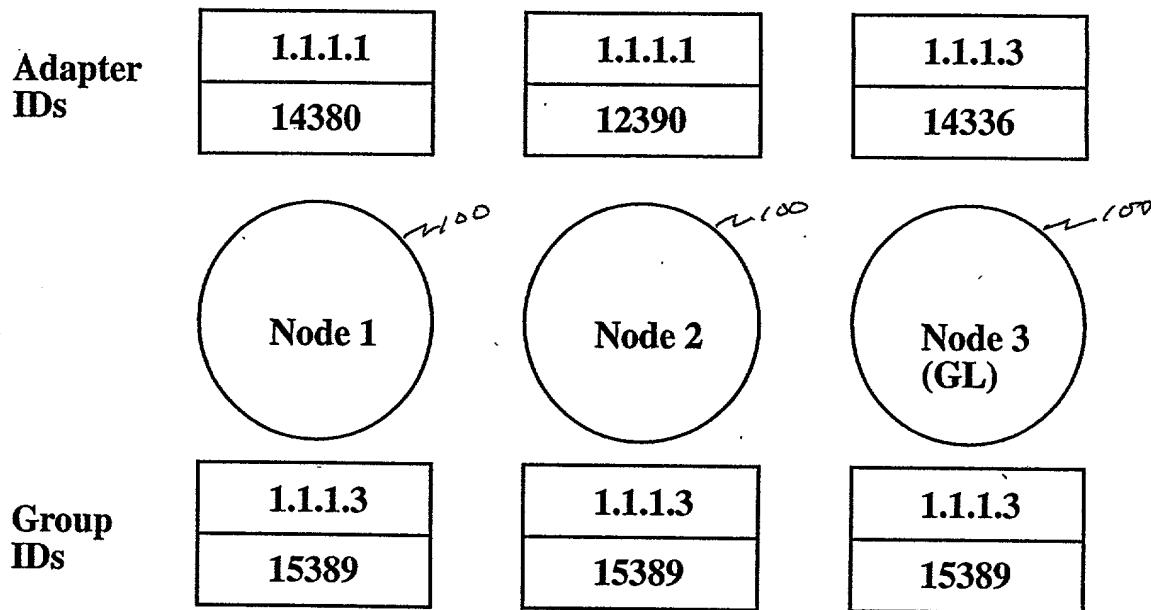
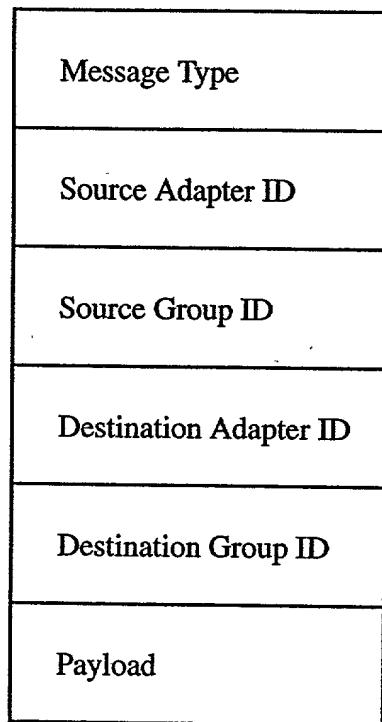


Figure 10) Adapter IDs and Group IDs. An adapter ID has the IP address of the adapter and and instance number. The Group ID has the IP address of the Group Leader and an instance number that changes each time the group changes.

17/30

Post 92002 0018 451



10  
Figure 10) Format of the protocol packets that are sent over the network

1830

Pou 920020018 US1

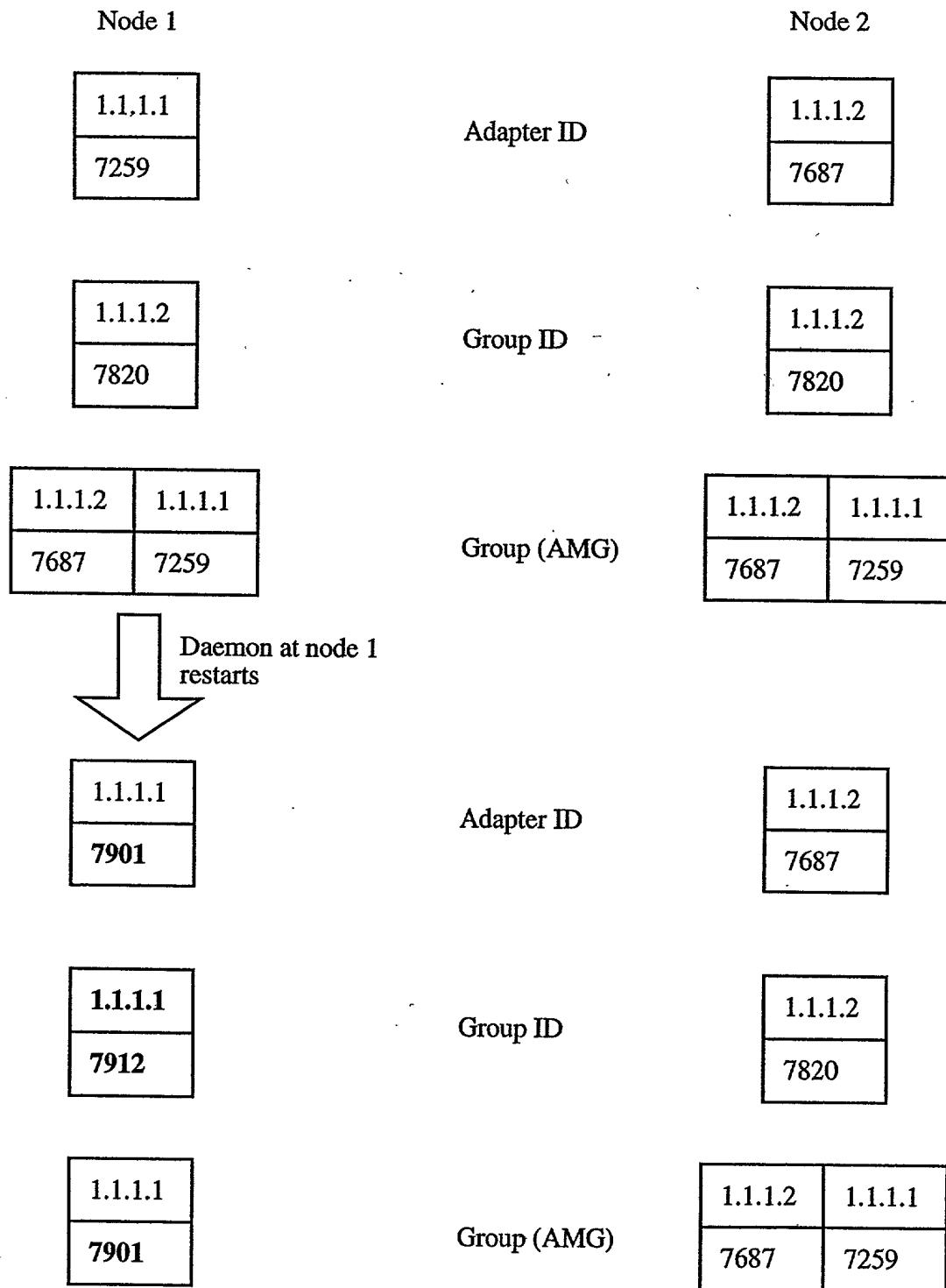
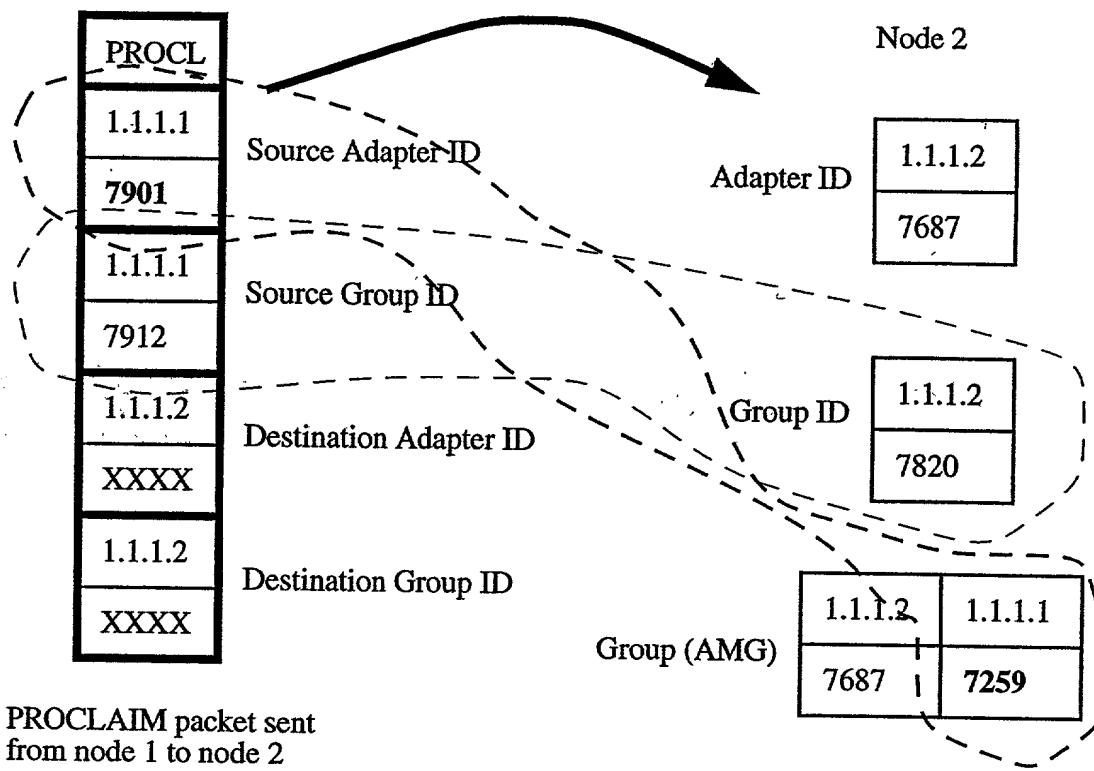


Figure 18) Adapter and Group IDs when the daemon at node 1 terminates and is restarted.

19/30

PD4920020018 US1



12  
Figure 12) A "live" node detects that a remote daemon restarted. The Group ID of the message is different from node 2's, while the address of the sender is listed on node 2's group membership.

20/30

Post 92002 0018 US1

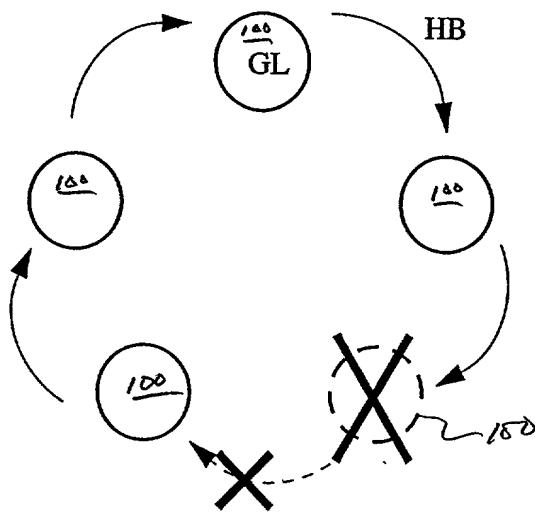
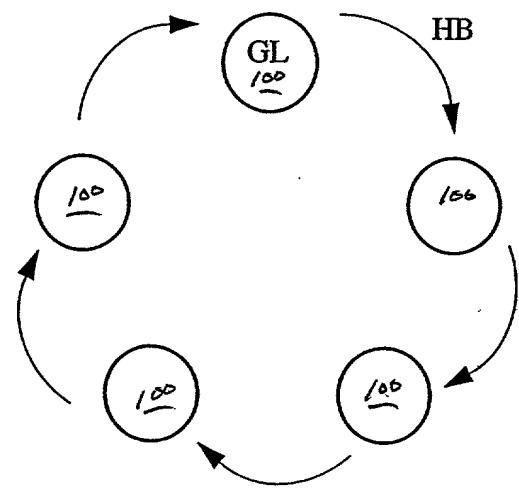
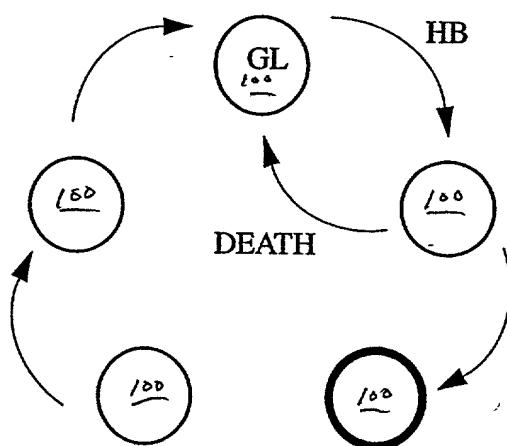
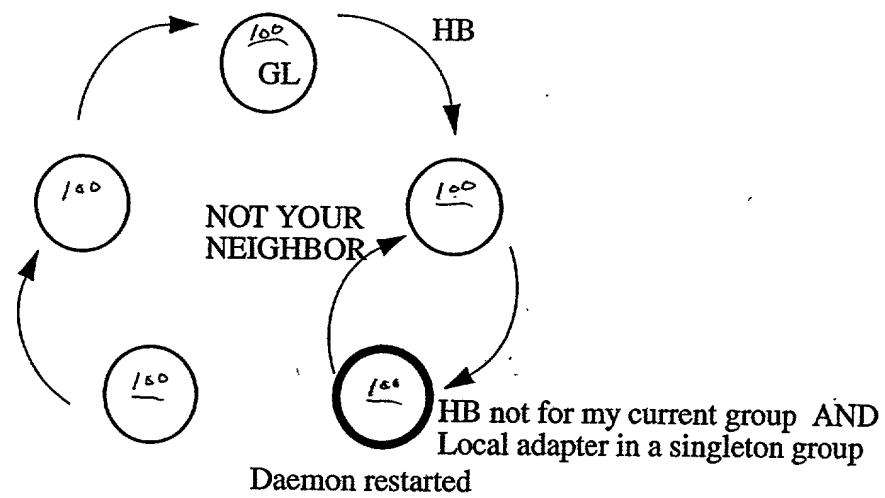


Figure 13a) A daemon that is restarted detects that a previous instance used to belong to an AMG because of heartbeat messages that it receives while in a singleton group.

21/30

P04920020018 US1



13  
Figure 2b) Continuation

22/30

P04920020018 US1

### Node 3 (GL)

1.1.1.3  
7687

### Adapter ID

Node 2

1.1.1.2  
7687

### Group ID

1.1.1.3  
7820

1.1.1.3

---

7820

1.1.1.3	1.1.1.2	1.1.1.1
7687	7259	7228

### Group (AMG)

1.1.1.3	1.1.1.2	1.1.1.1
7687	7259	7228

1.1.1.3	1.1.1.2	1.1.1.1
7687	7259	7228

### last\_stable\_group

1.1.1.3	1.1.1.2	1.1.1.1
7687	7259	7228

### Communication glitch in node 3's adapter.

Figure 5a) Solution to the Quick Communication Interruption Problem. initial state: nodes 1,2, and3 are part of the same AMG., Node 3's adapter suffers a temporary failure.

23/30

Post 9/2002 0018 451

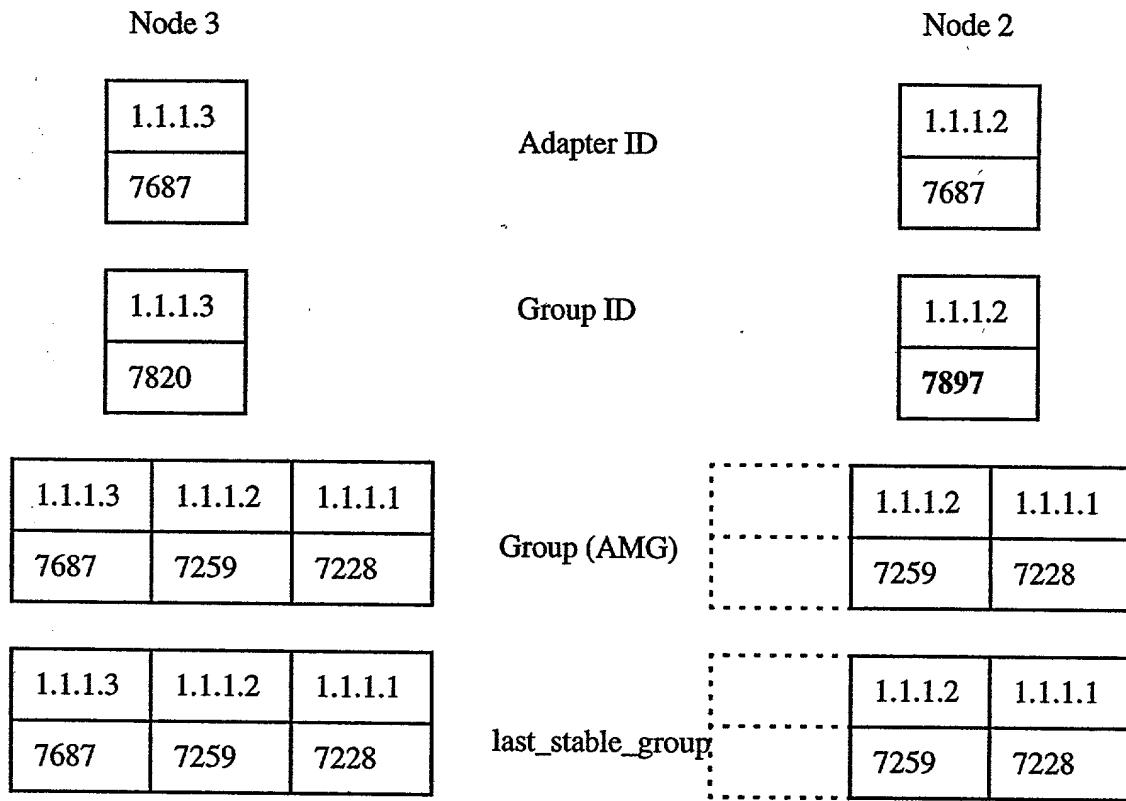


Figure 2b) Solution to the Quick Communication Interruption Problem. Node 3's adapter suffers a temporary failure. Node 2 commits a new AMG, while node 3 is still in the process of missing HBs from its neighbor

24/30

Pou920020018 251

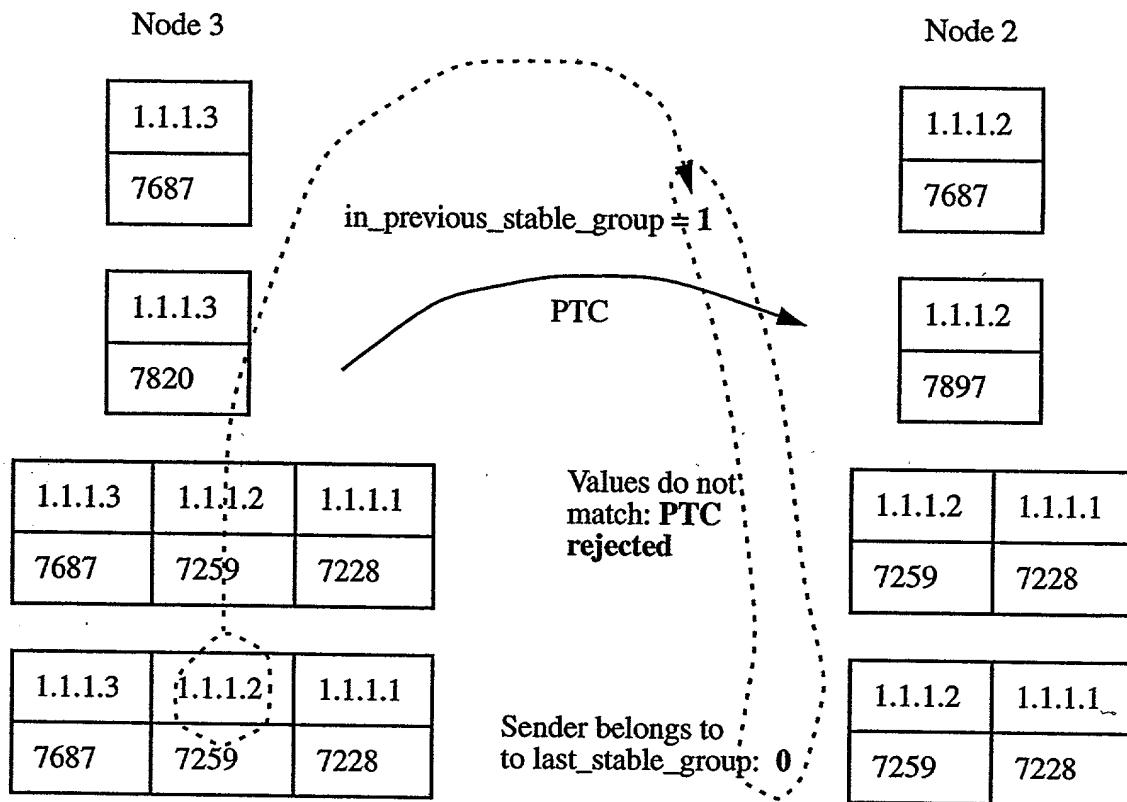


Figure 14  
c) Solution to the Quick Communication Interruption Problem. Node 3 sends a PTC when it stops receiving HBs from its upstream neighbor. The PTCs are rejected because of the discrepancy in the last\_stable\_group results.

25/30

Pou 920020018 451

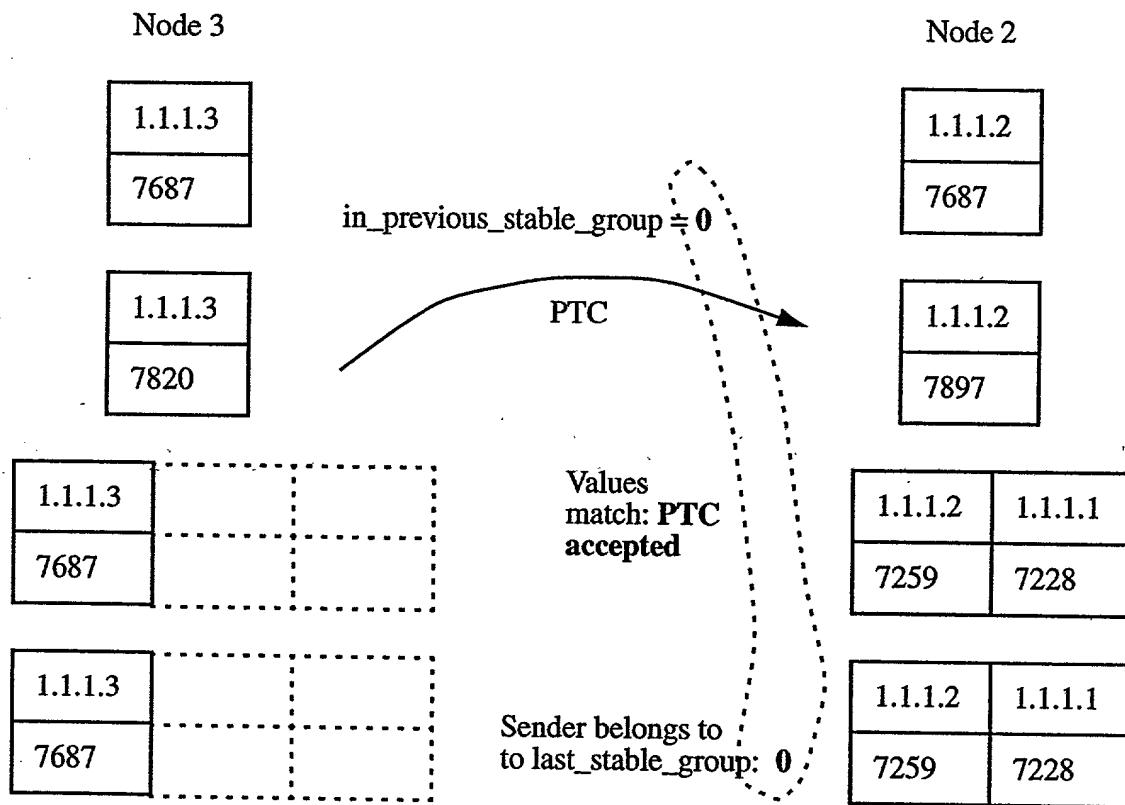
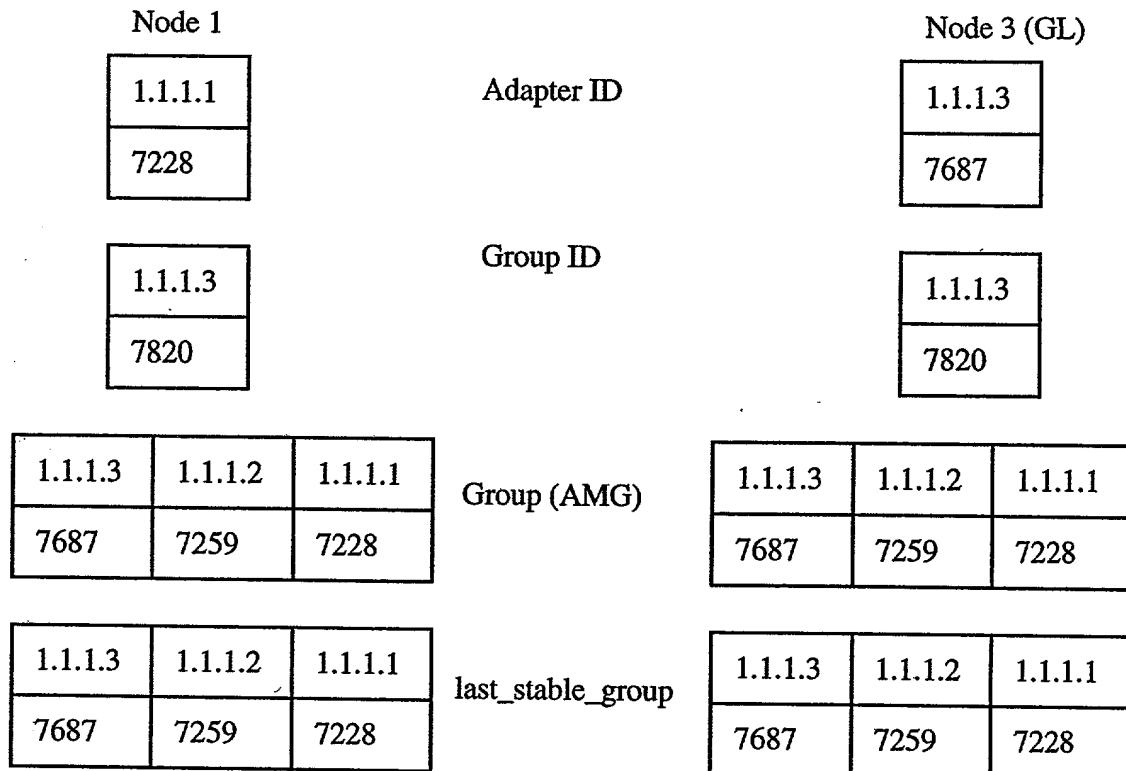


Figure 14d) Solution to the Quick Communication Interruption Problem. Since node 3 does not get replies to its PTC messages, it is forced to form a singleton group. At this point, it updates last\_stable\_group. From then on node 3's PTC are accepted again.

26/30

POU92002 0018 US1



Communication glitch  
in node 1's adapter.

Figure 2a) Solution to the Quick Communication Interruption Problem. initial state: nodes 1,2, and3 are part of the same AMG,. Node 1's adapter suffers a temporary failure.

27/30

Port 920020018451

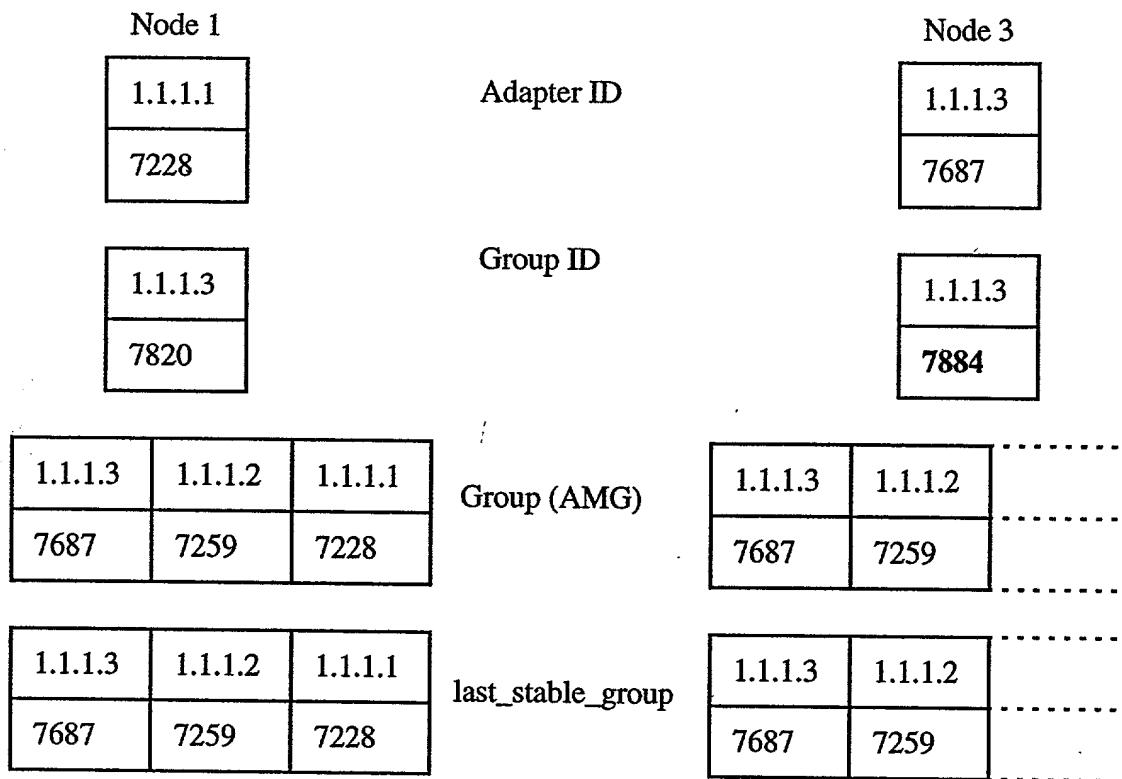


Figure 2b) Solution to the Quick Communication Interruption Problem. Node 1's adapter suffers a temporary failure. Node 3 commits a new AMG, while node 1 is still in the process of missing HBs from its neighbor

28/30

P0u920020018 US1

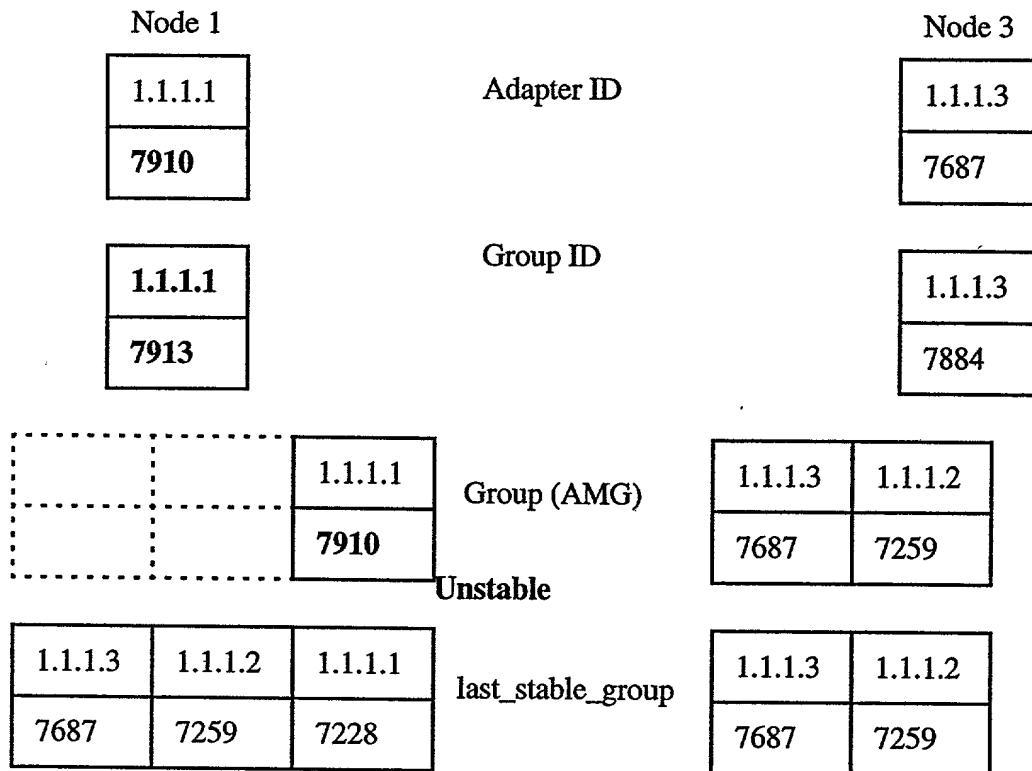


Figure 15) Solution to the Quick Communication Interruption Problem. Node 1 dissolves its group and forms a singleton unstable group. Note that because the group is unstable, there is no change in last\_stable\_group.

29/30

POU 920020018 USI

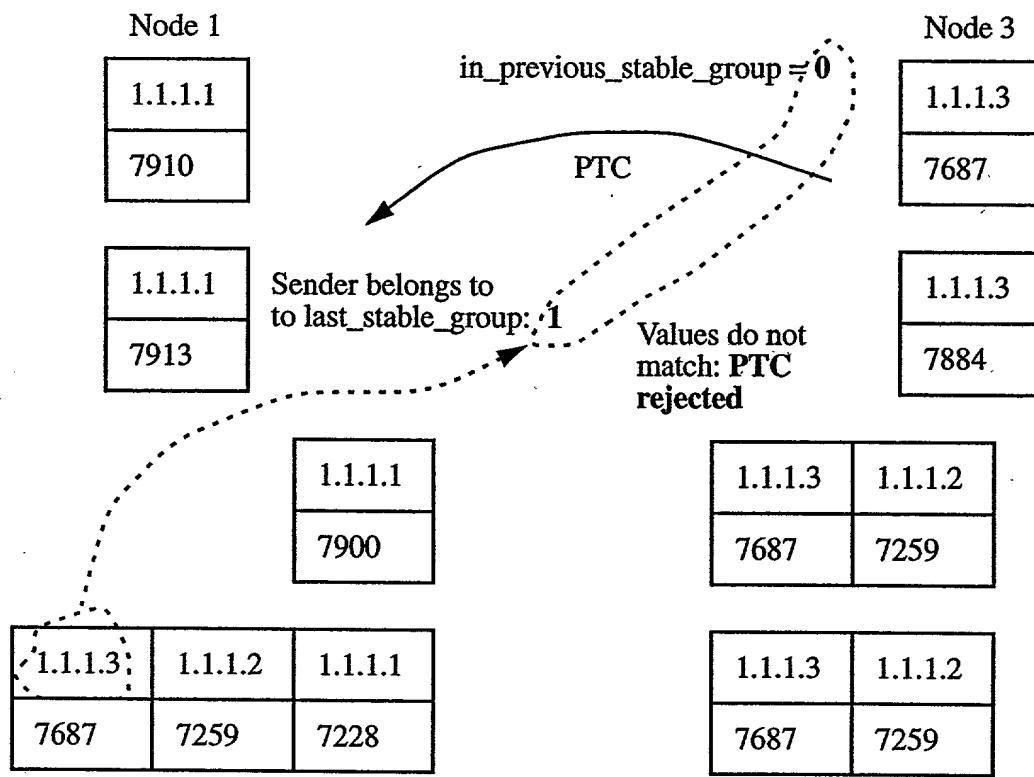


Figure 6(d) Solution to the Quick Communication Interruption Problem. Node 3 sends a PTC when node 1 responds the PROCLAIM message with a JOIN. The PTCs are rejected because of the discrepancy in the `last_stable_group` results.

30/30

Dom 9 2002 0018 451

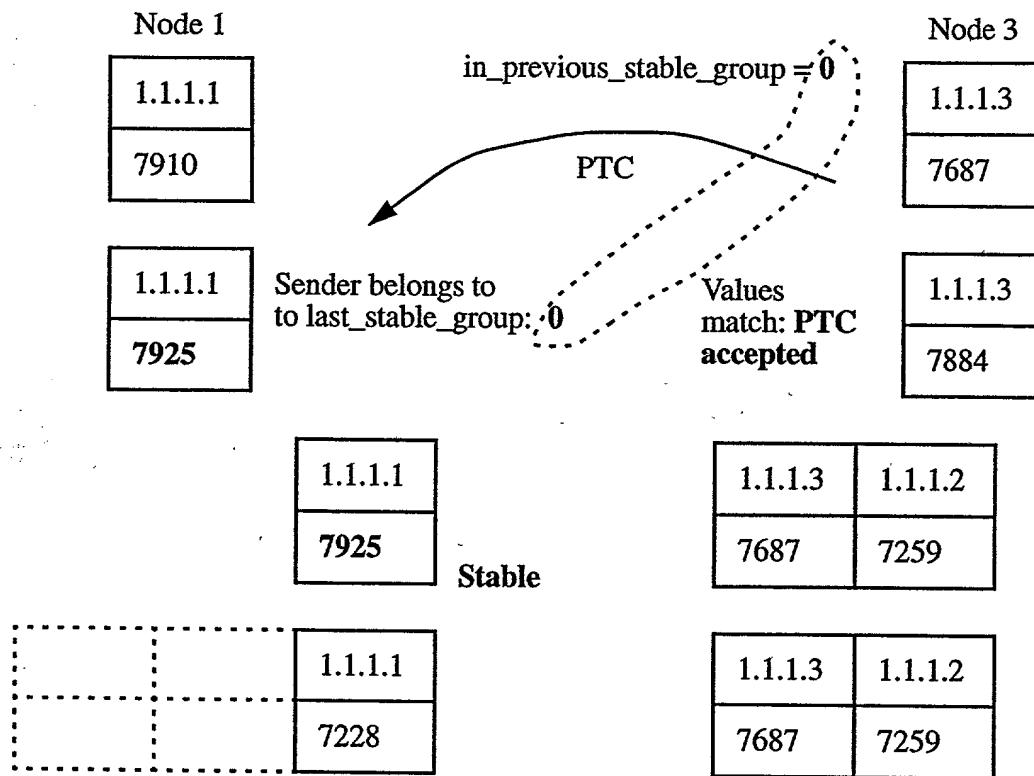


Figure 15e) Solution to the Quick Communication Interruption Problem. Since node 3 does not get replies to its PTC messages, it is eventually forced to form a singleton group. At this point, it updates last\_stable\_group. From then on node 3's PTC are accepted again.